

CLAIMS

What is claimed is:

1. A method comprising the computer-implemented steps of:
 - gathering statistics about XML resources that are stored in a database repository;
 - storing said statistics; and
 - in response to a request for access to one or more XML resources from said database repository, computing a computational cost associated with each of one or more methods of accessing said one or more XML resources from said database repository, based on said statistics.
2. The method of Claim 1, wherein said XML resources are logically organized in a hierarchy of nodes in which each node is either a container or a resource, and wherein the step of gathering statistics comprises gathering one or more data from a group consisting of
 - a total number of nodes in said hierarchy that are accessible via a path through a specified node,
 - a total number of containers in said hierarchy that are accessible via a path through said specified node,
 - a total number of nodes in said hierarchy that are accessible via a path through said specified node and that are in a level of said hierarchy that is immediately under a level of said specified node,
 - a total number of containers in said hierarchy that are accessible via a path through said specified node and that are in a level of said hierarchy that is immediately under said level of said specified node, and

15 a number of levels from a root node of said hierarchy, at which said specified
16 node is organized in said hierarchy.

1 3. The method of Claim 1, wherein said XML resources are logically organized in a
2 hierarchy of nodes in which each node is either a container or a resource, and
3 wherein the step of gathering statistics comprises gathering each of
4 a total number of nodes in said hierarchy that are accessible via a path through
5 a specified node,
6 a total number of containers in said hierarchy that are accessible via a path
7 through said specified node,
8 a total number of nodes in said hierarchy that are accessible via a path through
9 said specified node and that are in a level of said hierarchy that is
10 immediately under a level of said specified node,
11 a total number of containers in said hierarchy that are accessible via a path
12 through said specified node and that are in a level of said hierarchy
13 that is immediately under said level of said specified node, and
14 a number of levels from a root node of said hierarchy, at which said specified
15 node is organized in said hierarchy.

1 4. The method of Claim 1, wherein the step of storing statistics comprises storing said
2 statistics in a relational table of a database of which said database repository is part.

1 5. The method of Claim 4, wherein said relational table is a first relational table that is a
2 different table than a second relational table in which said XML resources are stored
3 in said database repository.

- 1 6. The method of Claim 4, wherein said relational table is a relational table in which
2 said XML resources are stored in said database repository.
- 1 7. The method of Claim 1, wherein the step of storing statistics comprises storing said
2 statistics in a hierarchical index table in which said XML resources are indexed to
3 said database repository.
- 1 8. The method of Claim 1, wherein the step of computing a computational cost
2 comprises computing a selectivity value for each of one or more predicates, from said
3 request, that contain operators on said database repository.
- 1 9. The method of Claim 8, wherein each of said XML resources is logically organized in
2 a hierarchy of nodes and stored, in association with a location of said XML resource
3 in said hierarchy, in a column of a table in said database repository, and wherein said
4 operator is an operator that determines whether a particular XML resource can be
5 located in said database repository through a particular specified path through a
6 portion of said hierarchy.
- 1 10. The method of Claim 8, wherein each of said XML resources is logically organized in
2 a hierarchy of nodes and stored, in association with a location of said XML resource
3 in said hierarchy, in a column of a table in said database repository, and wherein said
4 operator is an operator that determines whether a particular XML resource can be
5 located in said database repository at a terminal location of a particular specified path
6 through a portion of said hierarchy.

- 1 11. The method of Claim 1, wherein the step of computing a computational cost
2 comprises computing a computational cost of traversing, to locate a particular XML
3 resource specified in said request, an index in which said XML resources are indexed
4 to said database repository.
- 1 12. The method of Claim 11, wherein computing said computational cost of traversing an
2 index comprises computing a computational cost associated with one or more CPUs
3 used for said traversing.
- 1 13. The method of Claim 11, wherein computing said computational cost of traversing an
2 index comprises computing a computational cost associated with reading data blocks
3 in which portions of said index are stored.
- 1 14. The method of Claim 11, wherein computing said computational cost of traversing an
2 index comprises computing (a) a computational cost associated with one or more
3 CPUs used for said traversing and (b) a computational cost associated with reading
4 data blocks in which portions of said index are stored.
- 1 15. The method of Claim 1, wherein the step of computing a computational cost
2 comprises (a) computing a selectivity value for each of one or more predicates, from
3 said request, that contain operators on said database repository and (b) computing a
4 computational cost of traversing, to locate a particular XML resource specified in said
5 request, an index in which said XML resources are indexed to said database
6 repository.

- 1 16. The method of Claim 1, wherein said request for access to one or more XML
2 resources from said database repository is a SQL query.
- 1 17. The method of Claim 16, wherein each of said XML resources is logically organized
2 in a hierarchy of nodes and stored, in association with a location of said XML
3 resource in said hierarchy, in a column of a table in said database repository, and
4 wherein said SQL query comprises a mechanism for providing one possible path
5 through said hierarchy to each of said XML resources.
- 1 18. The method of Claim 17, wherein the step of computing a computational cost
2 comprises computing a computational cost component for one or more predicates,
3 from said request, that contain an operator in conjunction with said mechanism acting
4 on said database repository.
- 1 19. The method of Claim 16, wherein each of said XML resources is logically organized
2 in a hierarchy of nodes and stored, in association with a location of said XML
3 resource in said hierarchy, in a column of a table in said database repository, and
4 wherein said SQL query comprises a mechanism for providing all possible paths
5 through said hierarchy to each of said XML resources.
- 1 20. The method of Claim 19, wherein the step of computing a computational cost
2 comprises computing a computational cost component for one or more predicates,
3 from said request, that contain an operator in conjunction with said mechanism acting
4 on said database repository.

- 1 21. The method of Claim 1, wherein said database repository is part of a relational
2 database management system.
- 1 22. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 1.
- 1 23. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 2.
- 1 24. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 3.
- 1 25. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 4.
- 1 26. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 5.
- 1 27. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 6.

1 28. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 7.

1 29. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 8.

1 30. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 9.

1 31. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 10.

1 32. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 11.

1 33. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 12.

1 34. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 13.

1 35. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 14.

1 36. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 15.

1 37. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 16.

1 38. A method comprising the computer-implemented steps of:
2 gathering, by a database management system, statistics about XML resource that are
3 stored in a repository of said database management system; and
4 storing said statistics in said database management system.

1 39. The method of Claim 38, wherein the step of storing comprises storing said statistics
2 as XML data type in a schema-based table in said database management system.

1 40. The method of Claim 38, wherein said XML resources are logically organized in a
2 hierarchy of nodes in which each node is either a container or a resource, and
3 wherein the step of gathering statistics comprises gathering each of
4 a total number of nodes in said hierarchy that are accessible via a path through
5 a specified node,

6 a total number of containers in said hierarchy that are accessible via a path
7 through said specified node,

8 a total number of nodes in said hierarchy that are accessible via a path through
9 said specified node and that are in a level of said hierarchy that is
10 immediately under a level of said specified node, and

11 a total number of containers in said hierarchy that are accessible via a path
12 through said specified node and that are in a level of said hierarchy
13 that is immediately under said level of said specified node.

1 41. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 38.

1 42. A method comprising the computer-implemented steps of:
2 in response to a request for access to one or more XML resources from a database
3 repository within a database management system,
4 accessing, from said database management system, statistics about the
5 structure of a hierarchy in which said one or more XML resources are
6 logically organized; and
7 computing a computational cost associated with each of one or more methods
8 of accessing said one or more XML resources from said database
9 repository, based on said statistics.

1 43. The method of Claim 42, wherein the step of computing a computational cost
2 comprises computing a selectivity value for each of one or more predicates, from said
3 request, that contain operators on said database repository.

1 44. The method of Claim 42, wherein the step of computing a computational cost
2 comprises computing a computational cost of traversing, to locate particular XML
3 resources specified in said request, an index in which said XML resources are
4 indexed to said database repository.

1 45. The method of Claim 42, wherein the step of computing a computational cost
2 comprises (a) computing a selectivity value for each of one or more predicates, from
3 said request, that contain operators on said database repository and (b) computing a
4 computational cost of traversing, to locate a particular XML resource specified in said
5 request, an index in which said XML resources are indexed to said database
6 repository.

1 46. A computer-readable medium carrying one or more sequences of instructions which,
2 when executed by one or more processors, causes the one or more processors to
3 perform the method recited in Claim 42.

1 47. A database system comprising:
2 an XML data repository within a relational database management system; and
3 a query optimizer that receives a database query and, in response to said query,
4 formulates a query execution plan based on computational costs of access
5 paths associated with XML data stored in said repository, wherein said

6 computational costs are based on statistics about an organizational structure of
7 said XML data.

1 48. A system comprising:
2 means for gathering statistics about XML resources that are stored in a database
3 repository;
4 means for storing said statistics; and
5 means for computing, in response to a request for access to one or more XML
6 resources from said database repository and based on said statistics, a
7 computational cost associated with each of one or more methods of accessing
8 said one or more XML resources from said database repository.